

# BEA: BESZÉLT NYELVI ADATBÁZIS

Gósy Mária – Gyarmathy Dorottya – Horváth  
Viktória – Grácsi Tekla Etelka – Beke András  
– Neuberger Tilda – Nikléczy Péter<sup>1</sup>

## Előzmények

A magyar (és néhány más nyelv) hangzó változatának egyfajta első gyűjteménye ismereteink szerint Balassa József nevéhez fűződik a 20. század elején. Ez az értékes anyag azonban megsemmisült (l. Ismertetés a Keleti Kereskedelmi Akadémia Fonetikai Laboratóriumának munkájáról. A KKA 25. évi jelentése az 1915–16-iki iskolaév végén. Bp. 1916. 55–56. Közli: *Studia Academiae Nyíregyháziensis*. Tomus III., 1994). Az 1940-es években Hegedűs Lajos fonetikus kezdeményezésére indult el a nyelvjárási hangfelvételek készítése. Munkájuk során az ország különböző helyein rögzítettek beszédet, mesemondást, ráolvasásokat stb.; ezt az anyagot az MTA Nyelvtudományi Intézet archiváltatta (a 20. század kilencvenes éveinek végén), korszerű adathordozókra mentette (Gósy et al. 2011). A létrehozott adatbázisban 1200 hangfelvétel adatai találhatók.

A hetvenes évek elején Szende Tamás a spontán beszéd gyakorisági tényezőinek elemzéséhez négyféle spontánbeszéd-korpuszt használt fel (1973). Osztályteremben készültek felvételek, illetve más, különböző spontán beszédanyagokat rögzítettek a hetvenes évek második felében (Keszler 1983). A Budapesti Szociolingvisztikai Interjú (BUSZI) a múlt század nyolcvanas éveinek végén készült, magnetofonra felvett, beszélőnként 2-3 órás interjút (beleértve a spontán beszélgetést) tartalmaz (Kontra 1988; Váradí 2003). Lejegyzése, illetve kódolása is megtörtént (alapvetően helyesírásban, de a szövegek sajátos jeleket tartalmaznak egyes hangzások szemléltetésére).

A magyar dialektológiai kutatások hagyományosan a terepmunkán és a kérdőíves adatgyűjtésen alapulnak. Természetesen itt is nagy jelentőségűek a beszédfelvételek, mint például a Hegedűs archívuméi. A Magyar Nyelvjárási Hangoskönyv – MNyHk (lásd: <http://geolingua.elte.hu>) a Magyar Nyelvjárási Olvasókönyv (Hajdú–Kázmér 1974) szövegeiből ad közre válogatást az eredeti hangfelvételekkel együtt. A nyelvjárási anyagok informatizálását, elemzését és térképezését Vékás Domokos nyelv-

---

<sup>1</sup> A tanulmány írói valamennyien az MTA Nyelvtudományi Intézet Fonetikai Osztályának munkatársai, akik egyúttal a BEA fejlesztői is.

szeti technológiai fejlesztése, a BihalBocs (<http://www.bihalbocs.hu>) tette lehetővé (Vékás 1999; Vargha 2008). A magyar nyelvjárások atlaszának (MNYA.) szerves része az a mintegy 461 órányi hangfelvétel, amely az 1960 és 1964 között lezajlott ellenőrző gyűjtések során készült (25 szlovákiai és 327 magyarországi kutatóponton); spontánbeszéd-felvételeket is tartalmaznak (Balogh–Végh 1975).

A BABEL (<http://alpha.tmit.bme.hu/speech/hdbbabel.php>) az első olyan magyar felolvasásokból álló beszédatadbázis, amely nemzetközi szabvány alapján készült alapkutatási és beszédfelismerési célokra. Hatvan beszélő olvasott fel szövegeket (Vicsi–Vig 1998). Az MTBA magyar telefonbeszéd-adatbázis (<http://alpha.tmit.bme.hu/speech/hdbMTBAhun.php>) vezetékes és mobiltelefonról rögzített beszédgyűjtemény, amely a magyar beszédtechnológiai kutatások támogatására készült; ötszáz adatközlő 297 vezetékes, 203 mobiltelefonról rögzített felolvasásait tartalmazza (Vicsi et al. 2002; Vicsi 2010). Az adatbázis általános fonetikai, nyelvészeti kutatásokhoz szolgál alapul, és statisztikai feldolgozási módszereken alapuló személyfüggetlen gépi beszédfelismerők, dialógusrendszerek létrehozását teszi lehetővé.

Számos nyelven fejlesztettek beszédatadbázisokat különféle célokkal. Ezek nagysága, felvételi körülményei, átiratai, beszélői nagy különbözőségeket mutatnak (Gósy 2012). Létrejöttüket egyfelől az a fokozódó igény előzte meg, hogy a fonetikusok, nyelvészek a valós nyelvhasználat tanulmányozására törekedtek. Másfelől a beszédtechnológiai igények, elsősorban a mesterséges beszédfelismerés megoldása indokolta a nagy beszédanyagok szükségességét. A hatalmas adatmennyiségek tárolását és feldolgozhatóságát pedig a számítógépes és ahhoz kapcsolódó technológia lehetővé tette, és egyre jobban biztosítja. Csupán megemlítünk, természetesen a teljesség igénye nélkül, néhány jelentősebb beszédatadbázist, mint a British National Corpus, a London–Lund korpusz, a CallHome elnevezésű adatbázis, a Kiel-korpusz, a Switchboard, a Verbmobil, a CSJ (Corpus of Spontaneous Japanese), avagy európai nyelveket reprezentáló adatbázisok, például az EUROM1 és a BABEL (l. Gósy 2012). A BEA tehát egyike a világban fellelhető nagyméretű adatbázisoknak, egyes paraméterei – mint látni fogjuk – a legjelentősebbek közé emelik.

E tanulmány célja az MTA Nyelvtudományi Intézet Fonetikai Osztályán készülő BEszélt nyelvi Adatbázis (BEA) fejlesztésének, elemzési eredményeinek és kutathatóságának bemutatása. Ez a gyűjtemény multifunkcionális, és kiemelkedőnek tekinthető a beszélők számát, a protokollját, a rögzítés körülményeit és az átirásait tekintve. A fejlesztésnek mintegy a felénél tartunk, ez a mérőföldkő és az eddig elért eredmények jogosítanak fel bennünket a BEA bemutatására.

## A BEA fejlesztésének kezdetei

Az MTA Nyelvtudományi Intézet Fonetikai Osztályán 2007 őszén indult meg a BEA adatbázis elméleti tervezése, a szaktudományos igények felmérése, a meglévő adatbá-

zisok megismerése. Ezt követte a tartalmi kritériumok meghatározása, előzetes döntések a hangzó anyag lejegyzési stratégiáiról, valamint a munkamegosztásról. A felvételek készítése során a kezdetektől szigorúan figyelembe vettük a Nyelvtudományi Intézetben érvényes *Humán vizsgálatokon alapuló nyelvészeti kutatások etikai szabályozásában*<sup>2</sup> (2007. április 16) foglaltakat. Megtörtént az anonimizálás folyamatának kialakítása, ennek eredményeként az adatközlők kódokkal szerepelnek az adatbázisban, személyük azonosíthatatlan. Még az év őszén tizenöt próbafelvételt rögzítettünk, majd megvitattuk a felvételi tapasztalatokat és a próbalejegyzéseket. A távlati cél 500 Budapesten élő személy beszédének rögzítése, amelyben igyekszünk arányosítani a nők és a férfiak részvételét, valamint az életkor szerinti megoszlást. Az iskolázottság kiegyenlített reprezentálására nem törekszünk, de tendenciaszerűen figyelünk ennek az alakulására is.

## A BEA protokollja

Az adatbázis hangfelvételeinek készítése meghatározott protokoll szerint történik. Ez a protokoll tartalmazza a különféle rögzítendő beszédanyagokat, azok tartalmi sajátosságait, valamint a résztvevők számát. Az adatbázis többségében spontán beszédanyagokat tartalmaz, azokon kívül rögzítenek mondatismétléseket és felolvasásokat is. A protokoll hat részből áll, amelyek a következők: mondatismétlés, narratíva, véleménykifejtés, tartalomösszegzés, társalgás, felolvasás. Az alábbiakban tömören ismertetjük ezek tartalmát.

1. A mondatismétlés anyaga 25 egyszerű és összetett mondat. A tesztmondatok úgy lettek összeállítva, hogy változatos grammatikai szerkezeteket, szórendet és koartikulációs szabályokat tartalmazzanak, például: *A nagymama specialitása kétséget kizáróan a barackbefőtt.; Még nem lehet tudni, milyen súlyos a buszbaleset sérültjeinek állapota.* A mondatok átlagosan 8–12 szóból állnak, ezek döntő többségükben 3-4 szótagosak. A huszonöt mondat megismétlése nem igényel magas szintű kognitív működést, az adatközlőnek döntően a rövid távú memóriáját kell tudnia megfelelően működtetni.

2. A narratívák az adatközlő életéről, családjáról, munkájáról, hobbijáról szólnak, többé-kevésbé összefüggő monologikus szövegek. Az emberek általában könnyedebben beszélnek önmagukról, a munkájukról, a családjukról vagy a hobbijukról.

3. A véleménykifejtés (amely nagyrészt szintén narratíva) az interjúkészítő által megadott témának a véleményezése. Néhány a témák közül: házasság, illetve együttélés; eutanázia; közlekedés a fővárosban; sztárok/celebek alkohol- és drogproblémái; ittasan okozott autóbalesetek; megúszott büntetések; a magántulajdon védelme; új

---

<sup>2</sup> Az „1995. évi CXIX. törvény a kutatás és a közvetlen üzletszerzés célját szolgáló név- és lakcímadatok kezeléséről”, valamint az „1992. évi LXII. törvény a személyes adatok védelméről és a közérdekű adatok nyilvánosságáról” figyelembevételével. A hivatkozott törvények a Nyelvtudományi Intézet belső honlapján is hozzáférhetők.

adók, új szabályozások; BKV/MÁV-sztrájk; motoros balesetek; szervek felajánlása. Ez már bizonyos fokig nehezebb beszédfeladat, hiszen ekkor az adatközlőnek át kell gondolnia, mi a személyes véleménye a témáról, abból mennyit kíván elmondani, mit kíván elhallgatni, mi lehet az általános társadalmi vélekedés, ő személy szerint ahhoz képest hol helyezkedik el, stb. A témák azonban általánosak és hétköznapiak; konkrét ismereteket, háttértudást nem kíván az arról történő véleménykifejtés. Az interjúkészítő a témát az adatközlő életkorának és a narratíva alapján hallott érdeklődési körének megfelelően választja ki.

4. A tartalomösszegzés voltaképpen irányított spontán beszéd. Az adatközlő felvételről meghallgat egy-egy szöveget (amit női hanggal rögzítettek, átlagos beszédtempóban), és ezt követően rögtön a saját szavaival el kell mondania annak tartalmát. Az egyik szöveg egy rövid tudománynépszerűsítő cikk (174 szavas; 1 perc 37 mp tartamú), a másik egy anekdotaszerű történet (270 szavas; 2 perc 5 mp tartamú). Ennek a résznek a relatív nehézsége abból adódik, hogy a mindennapi kommunikációban ritkábban történik, hogy egy hallott történetet vagy ismeretterjesztő szöveget a beszélőnek a saját szavaival kell elmondania. Ez a rész teljesen monologikus, a kísérltetvezető nem szólal meg közben.

5. A társalgásban az adatközlőn és az interjúkészítőn kívül egy további személy vesz részt a beszélgetésben. A téma változó, az élet mindennapjaihoz kapcsolódik; ugyanazon adatközlő esetében mindig különbözik a véleménykifejtés témájától. A társalgás témáiból: karácsony, húsvét ünneplése; mobiltelefon kisgyermekeknek; új KRESZ és a biciklisek; halálbüntetés; dohányzási tilalom a szórakozóhelyeken; éjszakai élet, szórakozási lehetőségek Budapesten. Az interjúkészítő a témát ekkor is az adatközlő életkorának és érdeklődési körének figyelembevételével választja ki. A tapasztalatok alapján a társalgás az a része a protokollnak, amelynek során az adatközlők a legjobban feloldódnak. Itt ugyanis nemcsak ők beszélnek, hanem további két személy is, van idejük többet gondolkodni (szemben a narratívákkal), és a helyzetből adódóan jobban kikapcsolhatók a felvétel körülményei.

6. Az utolsó részben az adatközlő kétféle szöveget olvas fel; a hangos olvasás az általános iskolában elsajátított készség, az adatközlőknek nem okoz problémát. Az egyik a huszonöt, korábban ismételt mondat, a másik egy tudománynépszerűsítő cikk felolvasása, amely 291 szóból áll.

## Hangrögzítés

A BEA hangfelvételeinek legfontosabb követelménye, hogy jó jel-zaj viszonyú, magas dinamikájú, széles frekvenciaátvitelű és torzításmentes legyen, tehát a felvételi minőség megfeleljen a hanganyag későbbi feldolgozásához. Ennek érdekében az akusztikai és az elektronikai rendszer maximális minőségi követelményének megteremtését tartottuk elsődleges célnak. Akusztikai szempontból a zajszigetelt és visszhangmen-

tes helyiség használata, elektronikus szempontból pedig a megfelelő hangfelvő és hangrögzítő rendszer kialakítása volt a cél. Ezeket a feltételeket egy állandó felvételi helyiség biztosította.

A hangfelvétel helye a Fonetikai Osztályon már korábban kialakított saját tervezésű, zajszigetelt szoba (csendesített szoba), amely méréseink szerint megfelelt a fenti követelménynek. A szoba méretei (a hangszigetelő réteget nem számítva):  $340 \times 210 \times 300$  cm. A hangcsillapítás mértéke a külső környezethez képest 50 Hz-en 35 dB, 250 Hz fölött pedig  $\geq 65$  dB. A szoba belső terének fala – az utószög elkerülése érdekében –  $54 \times 54$  cm-es hangtörő felületekkel van kialakítva. A folyosói nyílászáró két, egymástól 30 cm távolságra lévő ajtó, amelyek külön-külön nyithatók. Mind a külső, mind a belső ajtó jó hangcsillapítású. A belső ajtó a Magyar Rádió által is használt, különlegesen kiképzett zajszigeteléssel van ellátva. A szobában felvett hanganyagok jel-zaj viszonyának értékei igazolják, hogy a helyiség alkalmas jó minőségű hangfelvételek készítésére.

Az elektronikai rendszer kialakításánál mind a felvő mikrofon(ok), mind a keverő erősítő minőségénél a maximumra törekedtünk. A felvételek Audio-technika AT4040 típusú kardoid kondenzátormikrofonnal készülnek. Ez a típus egyedülállóan alacsony zajt, széles dinamik tartományt és tiszta kimeneti jelet szolgáltat. A mikrofon Phonic MM102 típusú 2 csatornás phantom tápos analóg keverőn keresztül csatlakozik a számítógéphez. A rögzítés digitális, közvetlenül a számítógépre történik a GoldWave hangeditáló szoftverrel, 44,1 kHz-es mintavételezéssel. Tárolás: 16 bit, 86 kbyte/s, monó. A visszahallgató típusa Behringer Truth B2031 kétutas aktív stúdiómonitor.

Az ismertetett felvételi körülmények, illetőleg a csendesített szoba állandósága biztosítja, hogy a rögzített hanganyagok minden adatközlőnél megegyezzenek, és így összehasonlíthatók legyenek.

## Beszéd felvételek: spontaneitás és természetesség

A spontánbeszéd-korpuszokat gyakran éri az a vád, hogy eredeti céljuk ellenére a mesterséges rögzítési körülményekből fakadóan a valóságban nem a mindennapok verbális kommunikációját tükrözik. Ennek megítélésében azonban keveredni látszik a spontán beszéd és a természetes beszéd fogalma. Ez a „fogalomzavar” abból adódik, hogy többen szinonimaként használják a spontaneitás és a természetesség fogalmakat. A beszéd akkor spontán, ha a beszélő mindenféle előzetes tervezés és felkészülés nélkül az adott helyzetben fogalmazza meg gondolatait, rendeli hozzájuk a megfelelő nyelvi formát, és hangosítja meg a közléseket. A beszéd spontaneitása nem függ attól, hogy a beszéd narratíva-e, vagy társalgásban hangzik el. A spontaneitás független a beszélő aktuális lelkiállapotától, avagy az ismereteitől.

A szokatlan beszédhelyzet, illetve a beszéd rögzítés ténye azonban hatással lehet (általában negatívan) a beszéd természetességére. A mesterséges szituáció, a „mikrofon”

hatást gyakorolhat a beszéd folyamatosságára vagy akár a megszokott légzéstechnikára. Befolyásolhatja továbbá a lexikai válogatást (az adatközlő nyilvánvalóan jobban odafigyel arra, hogy milyen szavakat használ), valamint bizonyos mértékben a kiejtést is (Lindblom 1990: hiperejtés és hipoejtés). A természetességre hatással van a beszéd-téma, a kapcsolódó területek ismerete és az átadás módja is. A beszélő – többnyire önkéntelenül – alkalmazkodik a beszédpartnerhez, ez befolyásolhatja beszédének megformálását (Bata–Grácsi 2009), ami szintén összefüggést mutat a természetességgel. A beszélő személyisége nagymértékben meghatározza beszédének a természetességét. Van, akit jobban zavar egy adott kommunikációs helyzet, van, akit kevésbé vagy egyáltalán nem. Van, aki gyorsan figyelmen kívül hagyja a mesterséges helyzetet, van, akinek ehhez hosszabb idő szükséges.

A szorongás hatással van a spontán beszéd folyamatosságára és a beszéd természetességére, de nem a spontán voltára. A mikrofon zavaró tényező (Szende 1973), hiszen azt jelzi, hogy az elhangzott közlés visszavonhatatlan, korrigálhatatlan, megőrződik. A természetes beszédhelyzetekre ez nem jellemző, ezért erre az átlagos (nem színész, tévé- vagy rádióriporter) adatközlők nincsenek felkészülve, nem rendelkeznek kialakult stratégiákkal az ebből fakadó zavar ellensúlyozására. A legtöbb emberben szorongást kelt az a tudat, hogy felveszik és később különböző szempontokból elemzik majd a beszédét. Ilyenkor úgy érzik, hogy mind a beszéd tartalmi szempontját, mind a szóhasználat választékosságát tekintve bizonyos elvárásoknak kell megfelelniük. A beszédrengzés tényéből adódó zavar azonban nem egyformán érint mindenkit; akadnak, akikre ez nagyobb mértékben hat, míg mások (láthatólag) egyáltalán nem foglalkoznak vele, látszólag nem frusztrálja őket az adott helyzet. A mikrofon okozta esetleges beszédmódosulások az ő esetükben nehezen vagy egyáltalán nem ítéltetők meg. A nem feszélyezett (nem szorongó) viselkedés nyilvánvalóan kapcsolatban áll a természetes beszéddel; a szorongás mértéke és a beszéd természetessége közötti összefüggés azonban számos további kérdést vet fel. A mindennapi verbális kommunikáció során is adódhatnak olyan körülmények, amelyek a beszélőben szorongást keltenek (pl. alá-fölérendeltségi viszonyok).

Szende szerint (1973) egy ilyen helyzetben a beszélők törekszenek a nyelvi norma megközelítésére, sokkal inkább, mint a mindennapi megnyilatkozásokban. A megfigyelői paradoxonnak (Labov 1979) köszönhetően „minél jobban biztosítjuk [...] az adat-, illetőleg a szövegrengzés technikai föltételeit [...], a kísérleti effektus miatt annál inkább eltávolodnak az adatközlők a megszokott nyelvhasználatuktól, a kutatók pedig egyre távolabb kerülnek eredeti céljuk elérésétől” (Kiss 1995: 37).

Wolfson szerint a természetes beszéd egyenlő a beszélő szándékának megfelelő beszéddel (1976). Másként fogalmazva a beszéd akkor tekinthető természetesnek, ha a folyamat működtetése nem okoz feltűnő, észlelhető nehézséget a beszélő számára. Nusbaum és munkatársai a természetességet a beszéd multidimenzionális, szubjektív minőségeként tartják számon (1995). Kimondhatjuk, hogy a beszéd akkor természetes, ha megfelel a beszélő személyiségének, egyéniségének, szokásos beszédének. Nem



lehet azonban egzakt fonetikai paramétereket kapcsolni a természetes beszéd egyes jellemzőihez. Ugyanazt a történetet ugyanis különbözőképpen (más-más stílusban, szóhasználat, hangképzéssel, prozódiaival stb.) meséljük el a szüleinknek, idősebb vagy fiatalabb idegennek, rokonunknak, a „haveroknak”, a közeli barátunknak. A változó fonetikai paraméterek ellenére azonban senki sem mondaná azt, hogy a felsorolt helyzetek bármelyike ne felelne meg a természetesség kritériumainak.

## A BEA felvételeinek menete

A felvétel egyes részeinek sorrendjét tudatosan úgy alakítottuk ki, hogy minél jobban elősegítse az adatközlő gyors alkalmazkodását a felvételi körülményekhez, valamint a kezdeti szorongás mielőbbi leküzdését. Ennélfogva a beszélőnek van ideje „hozzászokni” a felvételi körülményekhez, az interjúkészítőhöz, így beszéde és viselkedése egy bizonyos idő elteltével (közel) természetessé válik. Akadhat azonban olyan adatközlő, akinél a cél érdekében a felvételi sorrend megváltoztatása szükséges, ez rugalmasan megtehető (az eddigi tapasztalatunk alapján igen ritkán fordult elő).

Egy felvétel átlagosan 50 percig tart, az adatközlők beszédkedve azonban nagymértékben változó (vannak, akik szűkszavúak, röviden fogalmaznak életrészben, mások hosszan fejtegetik a gondolataikat). A felvétel egyes részeinek időtartamát az adatközlő határozza meg, amíg van mondanivalója, addig beszélhet, ezt az interjúkészítő soha nem állítja le. Ilyen módon az egyes protokollrészek időtartama is jellemzi a beszélőt.

A felvételek menete a következő. Az interjúkészítő elsőként arra kéri az adatközlőt, hogy mondjon el a mikrofonba néhány mondatot; ezek rögzítése már abban a helyzetben történik, amelyben az egész felvétel zajlik. A kísérletvezető ezt a próbafelvételt a Praat programban (Boersma–Weenink 2011) azonnal ellenőrzi: megfelelőek-e a felvevő program beállításai a beszélő hangerejéhez képest. Ha mindent rendben talál, akkor kezdődik meg a tényleges felvétel.

Az interjúkészítő a mondatismétléssel kezdi a rögzítést. Egyenként felolvassa a 25 mondatot, amit az adatközlőnek mondatonként vissza kell mondania. Ha az adatközlő esetleg nem emlékszik pontosan az elhangzott mondatra, a felvételvezető annyiszor ismétli meg, ahányszor szükséges.

Ezután a két spontán narratíva következik. Az elsőben az adatközlőt a munkájáról, hobbijairól kérdezzük. Ennek célja, hogy a beszélő felkészülés nélkül, minél hosszabban beszéljen – az interjúkészítő csak akkor szólal meg, ha úgy ítéli meg, valamilyen kérdés, reagálás továbbviheti a narratívát. Az ismerős, előzetes felkészülést nem igénylő témák nem titkolt célja az esetleges szorongás oldása, illetőleg az, hogy amennyire lehetséges, elterelődjön a figyelem a mikrofonról. A második spontán narratívában az adatközlőt arra kérjük, hogy mondja el a véleményét egy a közvéleményt aktuálisan foglalkoztató közéleti, társadalmi történesről, témáról, vitáról. A beszélő

véleményéhez ekkor felhasználja az élettapasztalatát, egyes emlékeit is, és ezeket a személyes tapasztalatokat beleszövi a véleménykifejtésbe. A kísérletvezető is mindig képvisel valamilyen véleményt (jobbára az adatközlőével ellentéteset, hogy egyfajta vitára vagy meggyőzésre sarkalljon), ezért ez a rész (noha a cél továbbra is az, hogy az adatközlő beszéljen hosszabban) már jobban interakció jellegű, mint az előző.

A következő felvételrészben elsőként a tudományos ismeretterjesztő szöveget, majd a történelmi anekdotát hallgatja meg az adatközlő. Ez az első olyan „feladat”, amely esetleg szorongásra adhat okot, és amit az adatközlők zöme valóban feladatként él meg (ennek gyakran hangot is adva). Ekkorra azonban már elég idő eltelt ahhoz, hogy hozzászokjanak a felvételi körülményekhez, és azok már ne nagyon zavarják őket. Az eddigi tapasztalatok szerint az elhangzottak összefoglalása különösebb nehézséget senkinek nem jelent. Előfordul, hogy a tartalomösszegzéseket követően az adatközlő véleményezi a szövegeket, minősíti a szövegek nehézségét.

A társalgásban a harmadik résztvevő a Fonetikai Osztály egyik fiatal munkatársa. A témát vagy témákat az interjúkészítő indítja el, és bizonyos értelemben ő vezeti a beszélgetést, de a beszélőpartnerek is alakíthatják a témát, illetve azt, hogy milyen arányban vesznek részt a társalgásban.

A felolvasáskor az adatközlőnek lehetősége van előre átolvasnia a szöveget; vannak azonban, akik nem élnek ezzel, és azonnal megkezdik a felolvasást.

A felvétel rögzítését követően adminisztráció történik. Az adatközlő aláír egy beleegyező nyilatkozatot, amely lehetővé teszi a felvétel tudományos célú felhasználását az anonimitás biztosítása mellett. Fontos, hogy az adatközlő a felvétel után írja alá a nyilatkozatot, mert így módja van a felvételt követően azt meghallgatni, esetleg törlést kérni. Az interjúkészítő ezek után még kitöltet egy adatlapot is minden adatközlővel olyan információkat rögzítve, amelyek bizonyos kutatásokhoz fontosak lehetnek: életkor, magasság, testsúly, dohányzási szokások, esetleges beszédhiba. Az interjúkészítő a fejlesztésnek a jelen szakaszában 86%-ban ugyanaz a személy (nő), a hármas társalgásban a beszélőpartner 70%-ban női, 30%-ban férfibeszélő.

A felvétel előtt a beszélők egy része bevallja, hogy izgul (sokszor a testtartás, a gesztusok is tükrözik a szorongást). Ezek a külsődleges jelek azonban többnyire viszonylag gyorsan megszűnnek, az adatközlők hamar feloldódnak, aminek utólag hangot is adnak. Némelyszer kapunk olyan visszajelzést, hogy mennyire jó volt beszélgetni, a beszélőnek – saját maga számára is – meglepően gyorsan sikerült feloldódnia, belehelyezkednie a beszélgetésbe, a témába. Megfogalmazzák, hogy a kérdések, a társalgás, a beszélőpartnerek és az egész légkör olyan volt, hogy nagyon gyorsan elterelte a figyelmüket a hangrögzítés és az idegen, illetve mesterséges helyzet tényéről. Egy idős néni a felvétel megkezdése előtt arról panaszkodott, hogy a legnagyobb bánata, hogy nem tud senkivel beszélgetni, ezért nagy örömmel jött a felvételre.

Egy beszélőfelvétel hasonlóan bizalmi szituáció, mint például egy orvossal vagy egy pszichológussal történő beszélgetés. A közös munka csak akkor hatékony, ha a kommunikációs partnerek rövid idő alatt egymásra tudnak hangolódni, és kialakul



köztük ez a bizalmi viszony. Ez esetünkben azt is jelenti, hogy az adatközlő nem mérlegeli minden hangos megnyilatkozását, hanem a mindennapokban szokásos módján beszél. Ez pedig kétszeresen is fontos, egyfelől a beszéd természetességét növeli általában, másfelől specifikusan is abban az értelemben, hogy az adatközlő nem a feltételezett elvárások szerint beszél és fogalmaz, hanem a megszokott verbális rutinját követi.

Mindebben az interjúkészítőnek meghatározó szerepe van. Egyfajta intuitív, empátikus készséggel kell rendelkeznie ahhoz, hogy igen rövid idő alatt elnyerje a beszélő bizalmát, és mintegy „régi ismerősként” beszélgessen vele. Már a beszélgetés kezdetén fel kell tudnia mérni a beszélgetőpartner „szorongási szintjét”, és ennek megfelelően igyekeznie kell oldani azt. Ennek érdekében – mint említettük – lehetősége van például felcserélni az egyes részek sorrendjét. A túlságosan szorongó adatközlő könnyebben hozzászokik a felvételi helyzethez, gyorsabban feloldódik, ha a protokollt a háromfős társalgással indítjuk. Egy eleve oldott, kötetlenebb beszélgetésbe ugyanis könnyebben bekapcsolódik. Többször előfordult, hogy az ismeretlenként odaérkező adatközlő a felvétel után megkérdezte, hogy vajon voltak-e róla esetleg előzetes információk, mert az interjúkészítő olyan témákat hozott fel, amik őt nagyon foglalkoztatják mostanában. Egy másik esetben a beszélgetés során az adatközlő azt mondta, hogy bár tudja, hogy ezt a felvételt kutatók meg fogják hallgatni és elemzik, de akkor is elmondja, hogy ... És még egy példa. Megtörtént, hogy a felvetett kérdésre az adatközlőnek az eredeti véleménye a beszélgetés végére az ellenkezőjére változott. Ez pedig csak úgy lehetséges, hogy közben nem a felvételre figyelt, hanem teljesen lekötötte az adott téma, a véleménykülönbség, illetve a saját hozzáállásának a módosítása. Ezek a „vallomások”, vélemények, megjegyzések és tények megerősítik azt, hogy a felvételek valóban közel természetes beszédet tartalmaznak. A témák a beszélőkből gyakran érzelmeket is kiváltanak, ami nagyban hozzájárul ahhoz, hogy megfelelően legyenek a felvétel rögzítéséről. Nemegyszer előfordult például, hogy a véleménykifejtés hevében trágár kifejezés „csúszott ki” az adatközlő száján, és csak utána döbbsen rá, hogy ez egy hangfelvétel során történt: *...és akkor mondtam hogy hát mondom akkor most elmennek a ...ba így széttéptem a bocsánat a felvétel jaj jaj jaj jaj kiszaladt na mindegy; veszek egy kiló kenyeret és otthon megeszem az abált szalonnát és megiszom otthon egy pohár bort és akkor [akkor] cssesse meg az állam.*

A véleménykifejtés és a társalgás témája a mindenkor adatközlő valószínűsíthető személyes érdeklődéséhez igazodik. Az interjúkészítő igyekszik közfelháborodást kiváltó, illetve a közvéleményt foglalkoztató aktuális (nem politikai jellegű) eseményekről kérdezni az adatközlő véleményét. Ezekről a témákról ugyanis általában mindenkinek van tudomása és véleménye is, így könnyen és sokat képes beszélni róla. Az interjúkészítő figyelembe veszi a labovi elvet (1979, 1984), és végig „alárendelt” szerepben marad. A témafelvetéskor is tekintettel vagyunk az adatközlőre; ha a választott téma nem tetszik neki, láthatólag nem akar beszélni róla, akkor az interjúkészítő azonnal vált. Előfordul, hogy a beszélő a véleménykifejtés folyamán

elakad, megszakad a gondolatmenete, illetőleg az interjúkészítő úgy ítéli meg, hogy túl szűkszavúan nyilatkozott az adott témában. Ilyenkor – az adatközlő személyiségétől függően – szándékosan ellenkező nézőpontú kérdésekkel és véleményformálással igyekezik továbblendíteni a társalgást.

## A BEA adatbázis statisztikai adatai

A BEA adatbázis jelenleg 280 felvételt tartalmaz, ami 260 óra, 6 percnyi hanganyagot jelent. A legrövidebb felvétel 24 perc és 27 másodperc hosszú, míg a leghosszabb ennek az ötszöröse, 2 óra 24 perc 47 másodperc. Egy felvétel átlagosan 50 perc időtartamú. Adatközlőink egynyelvű, köznyelvet beszélő budapestiek, 163 nő és 117 férfi, átlagéletkoruk 39,27 év (20–90 év). A beszélők iskolai végzettségét tekintve a felsőfokú végzettségűek vannak többségben (54,67%), de szintén magas az érettségizettek aránya (41,52%). Az adatközlők kis százaléka (3,46%) rendelkezik csupán általános iskolai végzettséggel, ők jellemzően az idősebb korosztályból kerültek ki, amikor a tanulás lehetősége még nem adatott meg mindenkinek.

Foglalkozásukat tekintve beszélőink között találhatók általánosan ismert munkaköröket betöltők, például pedagógusok, gyógypedagógusok, orvosok, autószerelők, informatikusok, adminisztrátorok, színészek; de akadnak közöttük viszonylag ritkább foglalkozásokat űzők, például: cserépkályha-készítő, jelmeztervező, katolikus pap, evangélikus lelkész, orgonaépítő, pókerjátékos, forgatókönyvíró és sajtótitkár.

## Lejegyzések

Az elkészült felvételek lejegyzése folyamatosan készül a fejlesztés kezdetén kialakított lejegyzési kritériumok szerint. Az elsődleges lejegyzési tapasztalat eredményeként további annotálási formákat is alkalmazni kezdtünk. 163 felvétel lejegyzése .doc formátumban készült el, 179 hanganyag pedig .trs (Transcriber) fájlban érhető el. Több hanganyag mindkét módon elérhető. A Transcriberben készült lejegyzések .TextGrid (Praat) formátumba történő átalakítása folyamatban van. Ezek a .trs és .TextGrid címkeanyagok alapvetően a .doc lejegyzés szabályait követik az időbeni címke-hozzárendeléssel kibővülve a szoftver és a további kutatási célok függvényében némi módosítással. További, a Praat által támogatott formátumban egyéb lejegyzések is hozzáférhetők. A szakaszszintű címkék alapján automatikus hangszintű annotálás is készül (l. Gósy 2012).

A főbb lejegyzések jellemzőit részletesen ismertette Neuberger (2009, .doc), Gyarmathy és Neuberger (2011, .doc és .trs), illetve Gósy (2012). Ezért a jelen ismertetőben csak a lejegyzések közötti főbb eltérések irányelveit foglaljuk össze.

Mindhárom formátumban egy-egy lejegyző írja le a rögzített szövegeket egyéb, a kutatáshoz szükséges információk jelölésével.

A beszédanyag jelenleg legnagyobb arányban helyesírásban van lejegyezve, azonban a mondathatárokat nem jelölik. A mondatokra bontás elhagyásának oka a megnyilatkozáshatárok szubjektív és akár objektív megítélhetőségének nehézsége. Az irodalom ugyan felkínál kváziobjektív kritériumokat (Wacha 1988; Varga 1993; Németh T. 1996), ezek felhasználhatósága azonban nagymértékben relatív és szubjektív is, mert a részben egyöntetűnek tekinthető perceptuális jellemzők mellett gyakorta egyáltalán nem egységes a megnyilatkozás határainak megállapítása. A grammatikai, pragmatikai és prozódiai jellemzők felülírhatják, megkérdőjelezhetik a fogadó által valószínűsíthető szerkezetet, ennél fogva a határokat is (Gósy 2003; Váradi 2008; Markó 2010).

A lejegyzés bizonyos mértékben mindig szubjektív, hiszen az mindig az adott lejegyző saját percepciójának a függvénye (Hunston 2002). A jól betanított, jó adottságú lejegyzők relatíve kevés hibával rögzítik az elhangzottakat, de az ún. finom fonetikai jelenségek pontos tükröztetése lejegyzőnként eltéréseket mutathat. Az egy további lehallgató személy ellenőrzése ezen keveset segít, hiszen ő ugyancsak a saját beszédészlelési mechanizmusán keresztül dolgozza fel az elhangzottakat. Mindennek ellenére a lejegyzett hanganyagok nagyon nagy mértékben segítik a kutatásokat. A BEA-lejegyzések kapcsán Neuberger (2009) részletesen ír a szubjektív dekódolás sajátosságairól. Néhány jellemzőt emelünk ki ebből. A .doc formátumú lejegyzésekben a legközelebb eső magánhangzó háromszorozásával kell a hezitálásokat (kitöltött szüneteket) jelölni (pl. *eee, ööö*). A mindenkori lejegyző percepció szűrőjén is múlik, melyik magánhangzóval szemlélteti a hezitálást. A szünetek jelölése általában is szükségszerűen szubjektív, lejegyzőfüggő lehet. A Transcriberben végzett címkézések során a legalább 2 s-os szüneteket kell mindenképpen jelölni, míg a Wordben készült lejegyzésekben minden észlelt jelkimaradást egy-egy □ jel tükröztet. A szünetek észlelésével kapcsolatosan több kutatás igazolta, hogy számos tényezőtől függ (pl. Kassai 1988; Váradi 1988). A szünet helyzete (szerkezethatár, vagy sem), hossza, típusa (ti. kitöltött, kombinált, néma) és további jellemzők együttesen határozzák meg, hogy milyen valószínűséggel jelöli őket a hallgató, avagy hogy valójában nem megjelent szünetet létezőként észlel-e.

Az elhangzott közlések lejegyzése során felmerül egyes jelenségek jelölésének kérdése. Ilyen számos részletkérdést felvető probléma például a szünetek, a megakadások, a nonverbális vokális közlések, a beszélőváltások és az egyszerre beszélések megjelenítése. Ezekhez az egyes szoftverek lehetőségeiben elérhető jelölési módot kell alkalmazni, illetve az egyes jelenségek kapcsán a jelölések mélységét, finomságát a várható általános felhasználtság függvényében meghatározni. A három lejegyzés tehát az egyes szoftveres háttér jellemzőihez alkalmazkodva némi eltérést mutat, például a nem beszéd események jelölésében. Továbbá a lejegyzési szabályok egy-egy szempontot szükségszerűen jobban kiemelnek az egyik átiratban, mint a másikban,

a mindenkori kutatási, fejlesztési célhoz vagy várható főbb kutatási területekhez igazodva (lásd pl. Kontra 1988).

A lejegyzés szabályainak, szabályozottságának fontosságára például a zajok jelölését hozhatjuk fel. A .trs formátumú lejegyzésekben részletesebben különböztetjük meg az egyes zajokat, mint a .doc vagy a .TextGrid fájlokban. Az utóbbiakban a hangos légvétel, a köhögés, a száj- és nyelvcsettintés stb. !-lel, illetve !!!-lel jelenik meg, míg az előbbi esetében a megfelelő eseménycímkevel (breath, cough, lipsmack) jelölik az adatot a címsorban. A szoftver saját jellemzői a beszélőváltások jelölésében is megjelennek. A Word dokumentumban a beszélőváltást sortörés (és természetesen a beszélő azonosítójának megadása) jelzi, a Transcriberben új címke kezdete

T2 □ hát □ kertes ház

T1 ! és

T1 (\*bő\*)

T2 (de nem)

T2 inkább lakásban laknék és lemondok a kutyáról

T1 (!)

T2 (na mindegy)

T2 ! szóval szerintem kertes házba [házban] merhogy [merthogy] □ nekünk volt egy kutyánk ugyancsak de az szörnyű volt

T1 merhogy [merthogy] miért

T2 hát mer [mert] vonyított meg nyűszített meg hülye volt meg akit ő kitalált hogy az a főnök akkó [akkor] az a főnök aztán senki mást nem volt hajlandó elfogadni

T1 me [mert] nem nevelték meg rendesen

T2 én próbáltam de közölték hogy még kicsi aztán meg má [már] nevelhetetlen volt

T1 **hár** [hát] pedig kicsiként kell megnevelni

T2 ezt mondd el anyámnak

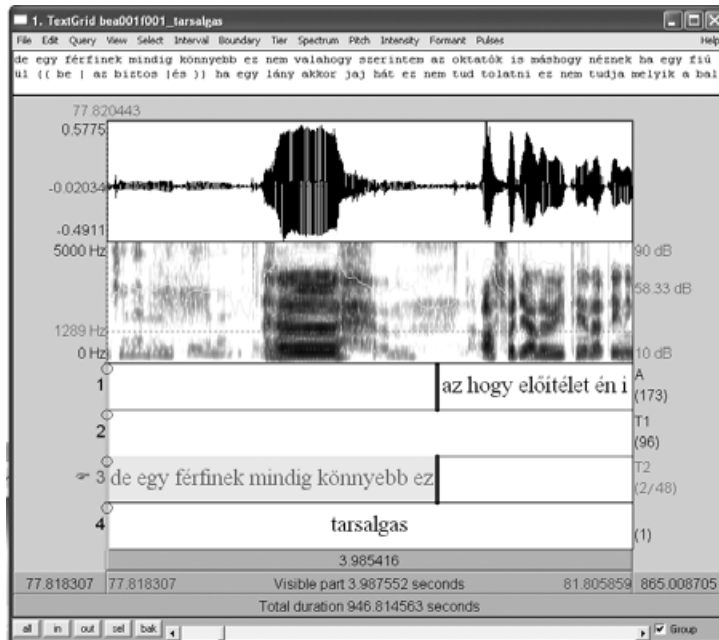
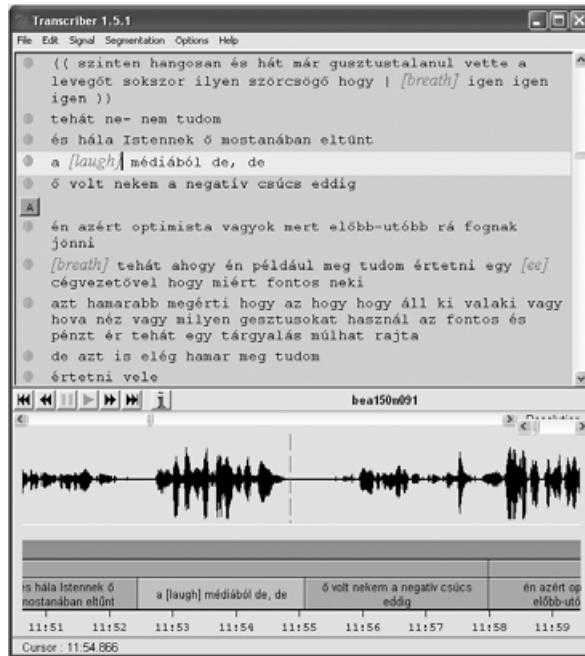
T1 ! jó

T2 (!)

T1 (! é-!)

T1 én az Arcsit [Archie-t] □ ő nekem egy kemény dió volt mert ő elég **őőő** ! erős természetű ellentétben az előző kutyámmal ! és őt például úgy kellett megnevelnem **hoogyő** semmi nem használt neki ami amúgy eddig bejött ! **me!** [mert] ő mindig **ki-** kikepezettte magát meg az ő akaratát akarta érvényre juttatni ! s **ztán** [aztán] végül ahhoz a trükkhöz folyamodtam **hoj** [hogy] szépen akkor ahogy a kutyák csinálják meg kell harapni kicsit így összecsipni a bőrt a nyakán ! és úgy lenyomni ! a földre más **más nemm é** nem jött be neki tehát **ezz** hiába **nyom-** feküdtem rá nyomtam le a földre

*1. ábra.* Példák a BEA adatbázis hanganyagainak lejegyzéseiből:  
egyszerű leírás megadott kritériumok alapján



1. ábra. Példák a BEA adatbázis hanganyagainak lejegyzéseiből: lejegyzés Transcriberben (fent) és Praatban (lent)

(és ugyancsak a beszélő azonosítójának megadása) jelöli; ezekkel szemben a Praat többszintű annotálást tesz lehetővé, amelynek előnye, hogy az egyes beszélők külön címkesorot kaptak, és az egyszerre beszéléseket leszámítva az egy beszélőhöz tartozó szöveg egy adott címkesorban jelenik meg (1. ábra). A párhuzamos beszédlépések, illetve beszédrészek jelölése az egyik részt vevő személynél jelenik meg. Ennek oka, hogy a Praatos címkesorokat a SLAAP projekt ([http://ncslaap.lib.ncsu.edu/tools/trans\\_to\\_praat.php](http://ncslaap.lib.ncsu.edu/tools/trans_to_praat.php)) által írt és ingyenesen hozzáférhető szkripttel konvertáljuk a Transcriber-fájlokból néhány szükséges jelölésbeli módosítás elvégzése után.

## A BEA beszédtechnológiai felhasználása

A beszédtechnológiában a beszédkorpusz számos területen alkalmazható, bizonyos problémák megoldása nem nélkülözheti a nagy mennyiségű rögzített beszédet. A BEA eleget tesz a beszédtechnológiai igényeknek, mivel jelentős mennyiségű beszélőt, minden beszélőtől hosszú és különböző stílusú, fonológiai, fonetikai, lexikai szempontból is igen gazdag spontán beszédmintákat tartalmaz, ezért alkalmas arra, hogy megfelelő akusztikai modelleket lehessen kinyerni belőle. Ezek az akusztikai modellek felhasználhatók például a beszédfelismerés számára, amelynek alapvető feladata, hogy a hangzó anyagot automatikusan szöveggé alakítsa.

A felolvasott beszéd szabályokkal relatíve jól megközelíthető, így annak gépi modellezése könnyebben megvalósítható, mint a spontán beszédé. Ezzel szemben a spontán beszéd sokkal kötetlenebb, nagyobb számú variációt tartalmaz, amelyek modellezése lényegesen összetettebb feladatot jelent. Az elmúlt évtizedben a spontán beszéd gépi feldolgozása, illetve modellezése egyre fontosabbá vált, hogy a humán, ember-ember kommunikációt minél pontosabban meg lehessen közelíteni. Ennek következményeként megnőtt az igény a spontánbeszéd-korpuszokra is.

A korpusz alkalmas arra is, hogy a beszéd-szintézishez kínáljon megfelelő akusztikai modelleket (HMM alapú beszéd-szintézis), illetve hogy a spontán beszédre igen jellemző megakadásjelenségeket – hezitálás, újraindítás, téves kezdések stb. – gépileg modellezni tudják, növelve ezzel a beszéd-szintézis természetességét.

Nagy mennyiségű beszédet tartalmazó adatbázis az igazságügyi hangtechnikai szakértés területén is felhasználható. Egyrészt a beszéd alapú személyazonosításban használatos akusztikai-fonetikai mérések referenciaanyagaként szolgálhat, másrészt része lehet az automatikus beszélőfelismerő rendszer populációs adatbázisának. Az automatikus beszélő felismerő rendszerek jellemzői közé tartozik a nyelv-, csatorna- és szövegfüggetlenség. Ennek a működését – többek között – a megfelelően sok beszélőt tartalmazó, felcímkézett adatbázis segíti elő. A BEA így lehetőséget biztosíthat a magyarul beszélő felnőttek mikrofonnal rögzített, spontán és olvasott beszédének összehasonlító hangszakértői vizsgálatára. Olyan algoritmusok kialakítására is



lehetőség van, amelyek képesek a beszélő nemét, illetőleg életkorát automatikusan osztályozni a beszédük alapján.

A napjainkban egyre fontosabb feladattá váló beszélődetektálás feladatának megvalósítására is használható az adatbázis, mivel nagyszámú társalgást (is) tartalmaz. A beszélődetektálás során a folyamatos beszédben az akusztikai jelből gépi úton határozzuk meg, hogy mikor ki beszél (Jin et al. 2004). A beszélődetektálást alapvetően két alfeladatra lehet bontani (Jin et al. 2004; Kotti et al. 2008): a beszélő szerinti szegmentálásra és a beszélőazonosításra. Az első feladat célja elkülöníteni az azonos beszélőktől származó beszédrészeket, a második részfolyamatban pedig ezeket a szegmentumokat kell osztályozni a beszélők szerint. Ennek során a folyamatos társalgások automatikusan lejegyzett szövegeit újrastrukturáljuk, így a szöveg sokkal könnyebben feldolgozható más, például tartalomkinyerő algoritmusok számára.

## Irodalom

- Bata Sarolta – Grácz Tekla Etelka 2009. Hatással van-e a beszédpartner életkora a beszélő beszédének szupraszegmentális jellegzetességeire. In Keszler Borbála – Tátrai Szilárd (szerk.): *Diskurzus a grammatikában, grammatika a diskurzusban*. Tinta Kiadó, Budapest, 74–83.
- Balogh Lajos – Végh József 1975. A magyar nyelvjárások atlaszához kapcsolódó hangfelvételek. In Deme László – Imre Samu (szerk.): *A magyar nyelvjárások atlaszának elméleti, módszertani kérdései*. Akadémiai Kiadó, Budapest, 257–262.
- Boersma, Paul – Weenink, David 2011. Praat: doing phonetics by computer (Version 5.1.34) [Computer program]. <http://www.praat.org> (a letöltés ideje: 2011. szeptember 2.)
- Gósy Mária 2003. Virtuális mondatok a spontán beszédben. *Beszéd kutatás 2003*. 19–43.
- Gósy Mária 2012. Multifunkcionális beszélt nyelvi adatbázis – BEA. *Általános Nyelvészeti Tanulmányok* 24. 329–349.
- Gósy Mária – Horváth Viktória – Nikléczy Péter 2011. A Hegedüs-archívum mint korszerű adatbázis. In Bárh M. János – Vargha Fruzsina Sára (szerk.): *Hangok – helyek*. ELTE Magyar Nyelvtudományi és Finnugor Intézet, Budapest, 85–103.
- Gyarmathy Dorottya – Neuberger Tilda 2011. A BEA adatbázis alkalmazásfüggő lejegyzései. *Beszéd kutatás 2011*. 109–121
- Hajdú Mihály – Kázmér Miklós 1974. *Magyar nyelvjárási olvasókönyv*. Tankönyvkiadó, Budapest.
- Hunston, Susan 2002. *Corpora in applied linguistics*. Cambridge University Press, Cambridge.
- Jin, Qin – Laskowski, Kornel – Schultz, Tanja – Waibel, Alex 2004. Speaker segmentation and clustering in meetings. In *Proceedings of ICSLP*. Jeju, Korea.
- Kassai Ilona 1988. A szünet kérdésköre a szöveglejegyzésben. In Kontra Miklós (szerk.): *Beszélt nyelvi tanulmányok*. Linguistica, Series A, Studia et Dissertationes 1. MTA Nyelvtudományi Intézet, Budapest. 22–43.
- Keszler Borbála 1983. Kötetlen beszélgetések mondat- és szövegtani vizsgálata. In Rácz Endre – Szathmári István (szerk.): *Tanulmányok a mai magyar nyelv szövegtana köréből*. Akadémiai Kiadó, Budapest, 164–202.
- Kiss Jenő 1995. *Társadalom és nyelvhasználat*. Nemzeti Tankönyvkiadó, Budapest.

- Kontra Miklós 1988. Bevezető. In Kontra Miklós (szerk.): *Beszélt nyelvi tanulmányok*. Linguistica, Series A, Studia et Dissertationes 1. MTA Nyelvtudományi Intézet, Budapest, 1–4.
- Kotti, Margarita – Benetos, Emmanouil – Kotropoulos, Costas 2008. Computationally efficient and robust BIC-based speaker segmentation. In *IEEE Trans. Audio, Speech, Language Processing*. Vol. 16/5. 920–933.
- Labov, William 1979. A nyelv vizsgálata társadalmi összefüggésben. In Pléh Csaba – Terestyéni Tamás (szerk.): *Beszédaktus – kommunikáció – interakció*. Tömegkommunikációs Kutatóközpont, Budapest, 365–398.
- Labov, William 1984. Field methods of the project on linguistic change and variation. In Baugh, John – Sherzer, Joel (eds.): *Language in use: Readings in sociolinguistics*. Englewood Cliffs, NJ, Prentice-Hall, 28–53.
- Lindblom, Björn 1990. Explaining phonetic variation: a sketch of the H & H theory. In Hardcastle, William J. – Marchal, Alain (eds.): *Speech production and speech modelling*. Kluwer Academic Publishers, Dordrecht, 403–439.
- Markó Alexandra 2010. A prozódia szerepe a spontán beszéd tagolásában. *Beszédkutatás* 2010. 82–99.
- Németh T. Enikő 1996. *A szóbeli diskurzusok megnyilatkozáspéldányokra tagolása*. Nyelvtudományi Értekezések 142. Akadémiai Kiadó, Budapest.
- Neuberger Tilda 2009. A spontán beszéd lejegyzése – a BEA adatbázis tapasztalatai alapján. *Beszédkutatás* 2009. 182–195.
- Nusbaum, Howard C. – Francis Alexander L. – Henly, Anne S. 1995. Measuring the naturalness of synthetic. *International Journal of Speech Technology* 1. 7–19.
- Szende Tamás 1973. *Spontán beszédanyag gyakorisági mutatói*. Nyelvtudományi Értekezések 81. Akadémiai Kiadó, Budapest.
- Váradai Tamás 2003. A Budapesti Szociolingvisztikai Interjú. In Kiefer Ferenc – Siptár Péter (szerk.): *A magyar nyelv kézikönyve*. Akadémiai Kiadó, Budapest, 339–359.
- Váradi Viola 2008. A virtuális mondatok műfaji meghatározottsága. *Beszédkutatás* 2008. 134–147.
- Varga László 1993. *A magyar beszéddallamok fonológiai, szemantikai és szintaktikai vonatkozásai*. Nyelvtudományi Értekezések 153. Akadémiai Kiadó, Budapest.
- Vargha Fruzsina 2008. Nyelvjárási és helynévtörténeti anyagok számítógépes feldolgozása. In Alabán, František (szerk.): *Kontextus – Filológia – Kultúra II*. Banská Bystrica, 75–82.
- Vékás Domokos 1999. Informatikai lehetőségek a dialektológiában, különös tekintettel a fonetikai szempontokra. Kézirat. ELTE Fonetikai Tanszék, Budapest.
- Vicsi Klára 2010. Adatbázisok a beszédtechnológia szolgálatában. In Német Géza – Olasz Gábor (szerk.): *A magyar beszéd – beszédkutatás, beszédtechnológia, beszédinformációs rendszerek*. Akadémiai Kiadó, Budapest, 262–332.
- Vicsi Klára – Vig Attila 1998. Az első magyar nyelvű beszédadatbázis. *Beszédkutatás* '98. 163–178.
- Vicsi Klára – Tóth László – Kocsor András – Gordos Géza – Csirik József 2002. MTBA – magyar nyelvű telefonbeszéd-adatbázis. *Híradástechnika* 8. 35–39.
- Wacha Imre 1988. Élő nyelvi (spontán) szövegek megnyilatkozásainak (szintaktikai) vizsgálati szempontjaihoz (a gazdagréti kábeltelevízió élő nyelvi felvételei alapján). In Kontra Miklós (szerk.): *Beszélt nyelvi tanulmányok*. Linguistica, Series A, Studia et Dissertationes 1. MTA Nyelvtudományi Intézet, Budapest, 102–158.
- Wolfson, Nessa 1976. *Speech events and natural speech: Some implications for sociolinguistic methodology*. Cambridge University Press, Cambridge.